Study and characterization of odd and even nonlinearities in electrodynamic loudspeakers by means of periodic random-phase multisines

Gil-cacho, Pepe; Van Watershoot, Toon; Moonon, Marc; Jensen, Søren Holdt

Published in:
Proceedings of the 127 AES Convention

Publication date:
2009

Document Version
Publisher's PDF, also known as Version of record

Audio Engineering Society

Convention Paper 7841

Presented at the 127th Convention
2009 October 9–12 New York NY, USA

Study and characterization of odd and even nonlinearities in electrodynamic loudspeakers

Pepe Gil-Cacho[1], Toon van Waterschoot[1], Marc Moonen[1], and Søren Holdt Jensen[2]

[1]*Katholieke Universiteit Leuven, ESAT-SCD, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium*

[2]*Aalborg University, Dept. Electronic Systems, Niels Jernes Vej 12, DK-9220 Aalborg, Denmark*

Correspondence should be addressed to Pepe Gil-Cacho (`pepe@esat.kuleuven.be`)

ABSTRACT

In acoustic echo cancellation (AEC) applications, often times an acoustic path from a loudspeaker to a microphone is estimated by means of a linear adaptive filter. However, loudspeakers introduce nonlinear distortions which may strongly degrade the adaptive filter performance, thus nonlinear filters have to be considered. In this paper, measurements of three types of loudspeakers are conducted to detect, quantify and qualify nonlinearities by means of periodic random-phase multisines. It is shown that odd nonlinearities are more predominant than even nonlinearities over the entire frequency range. The aim of this paper is then to demonstrate that third-order (cubic) adaptive filters have to be used, which is in clear conflict with the extensive, almost unique, use of second-order (quadratic) Volterra filters.

1. INTRODUCTION

Acoustic echo cancellation (AEC) is used in speech communication applications where the existence of echoes degrades the intelligibility and listening comfort, such as in mobile and hands-free telephony and in teleconferencing. The general set-up of an acoustic echo canceller is depicted in Fig.1. An acoustic canceller seeks to cancel the echo signal component $y(t)$ in the microphone signal $d(t)$, ideally leading to an *echo-free* error signal $e(t)$ which is then sent to the far-end side. This is done by subtracting an estimate of the echo signal $\hat{y}(t)$ from the microphone signal. The echo signal is the far-end signal $u(t)$ filtered by the loudspeaker-enclosure-microphone (LEM) impulse responses or echo path. Therefore, an adaptive filter is used to provide a model $\hat{G}$ that represents the best fit to the echo path [3]. This model is used to filter the far-end signal $u(t)$ to obtain the estimated echo signal. Standard approaches to AEC rely on the assumption that the echo path can be modelled by a linear filter. While the room (enclosure) and microphone responses may be considered as linear, the loudspeaker often introduces nonlinear distortions and hence it must be modelled as a nonlinear system. A nonlinear sytem will transfer energy from one frequency to other frequencies, and so the response will contain combinations of the input harmonics. The response of an even nonlinearity (e.g., x^2) to even and odd harmonics contains, in both cases, even harmonics. The response of an odd nonlinearity (e.g., x^3) depends on the harmonic content of the input signal, so the response to even and odd harmonics contains, respectively, even and odd harmonics only.

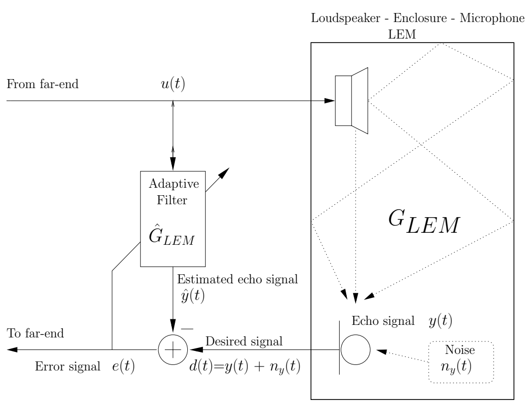

Fig. 1: Acoustic echo canceller set-up

In [5] it is shown that for random excitations, a nonlinear system can be represented by a linear system followed by an additive noise source (see Fig.2). For the considered class of excitation signals, the linear system provides the best linear approximation (BLA) of the output signal. The noise source represents the nonlinear distortions not represented by the linear system.

The frequency response function (FRF) $G(jw_k)$ of the nonlinear system may be written as

$$\begin{aligned} G(jw_k) &= G_{BLA}(jw_k) + G_s(jw_k) \\ &= G_0(jw_k) + G_B(jw_k) + G_s(jw_k) \end{aligned} \quad (1)$$

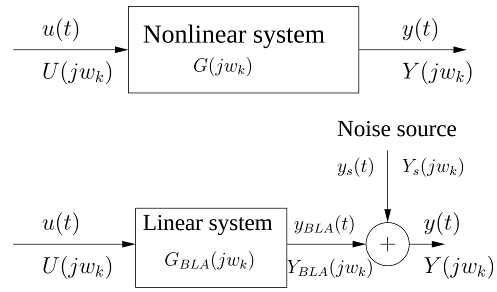

Fig. 2: Representation of a nonlinear system as the combination of a linear system plus a noise source

Each term in expression (1) contributes to the echo path in different ways:

$G_0(jw_k)$ is the true underlying linear system and it forms together with $G_B(jw)$ the BLA.

$G_B(jw_k)$ is the systematic nonlinear contribution and it acts as a bias term on the estimated model. Only odd nonlinearities contribute to this term [5].

$G_s(jw_k)$ is the stochastic nonlinear contribution and it acts as a noise source. This is represented in Fig.2 where $Y_S(jw_k) = G_S(jw_k)U(jw_k)$. Odd and even nonlinearities contribute to this term [5].

Considering the output of a nonlinear system as $Y(jw_k) = Y_{BLA}(jw_k) + Y_S(jw_k)$ it can be seen that nonlinearities in the echo path have two important effects. First, the error signal $e(t)$ will be contaminated by noise (even in the absence of background noise $N_Y(jw_k)$) since the linear adative filter cannot model the nonlinear contribution in the echo path. As a consequence, the adaptive filter parameters will converge slowly or even diverge [3]. Second, the filter parameters will not converge to the true linear parameters due to the bias term in the best linear approximation [5]. Two questions may arise at this point: do the nonlinearities carry sufficient relevance such that nonlinear models have to be considered? If so, and in order to choose a proper nonlinear model, what types of nonlinearities are present in the system? In the AEC literature ([1],[2] and references therein) second-order distortions are commonly assumed and so, second-order Volterra filters are used extensively. In this paper, this assumption is checked experimentally, and it is shown that in fact odd (i.e., third-order) nonlinearities are more predominant than even (i.e., second-order) nonlinearities. The paper is organized as follows: In Section 2 a special class of excitation signals called random-phase multisines is presented, which will be employed throughout the paper. Then the measurement methods and the information extracted from these are explained. Section 3 presents the actual loudspeaker measurements, showing the level and type of nonlinearities in loudspeakers of different quality. Conclusions are drawn in Section 4.

2. DETECTION OF NONLINEAR DISTORTIONS

The presence of nonlinear distortions should be detected, quantified (i.e., estimating the level of the distortions) and qualified (i.e., classifying in even or odd distortions). To achieve this, a special class of signals, called periodic random-phase multisines, are used here. The use of this class of signals allows for separating the nonlinear distortion and the disturbing noise level [6] in a simple manner. This is done by analyzing the FRF amplitude variations over consecutive periods and over different phase realizations of the input, as will be explained. The analysis is performed assuming just output (i.e. background) noise $n_y(t)$. This is a very common assumption in AEC applications since the input (i.e. excitation) signals are the *.wav* files stored in the PC and therefore noiseless.

2.1. Excitation signals

A signal $u(t)$ is a random-phase multisine [5] excitation if

$$u(t) = \sum_{k=1}^{F} U_k \cos\left(2\pi k f_0 t + \phi_k\right) \tag{2}$$

with f_0 the frequency of the first harmonic which sets the resolution of the Discrete Fourier Transform (DFT) spectrum, F is the number of harmonics and U_k the deterministic amplitude of the k^{th} harmonic. The phases ϕ_k are realizations of independent uniformly distributed random processes on [0,2π) such that $E\left\{e^{\phi_k}\right\} = 0$. These signals combine the advantage of random behaviour and periodicity while the user has total control over the amplitude of each harmonic. The main advantage for our purpose is that these signals allow to obtain insight in the presence of nonlinear distortions.

2.2. Estimating the level of the stochastic nonlinear contributions[1]

To estimate the level of the stochastic nonlinear contributions $G_s(jw_k)$ several phase realizations of a full multisine are used. A full multisine is a signal as in (2) where harmonics $k_i \in \{k \mid k = 1,2,...,F\}$ (i.e., even and odd harmonics) are excited. This will allow to have the response to odd and even harmonics together, and will provide a complete picture of the total level of the stochastic nonlinear distortions. The measurement starts from FRF data $G(jw_k) = \frac{Y(jw_k)}{U(jw_k)}$ where $w_k = 2\pi k f_0$. The level of the stochastic nonlinear distortion is obtained through averaging over different random phase realizations of the multisine excitation as shown in Fig.3.

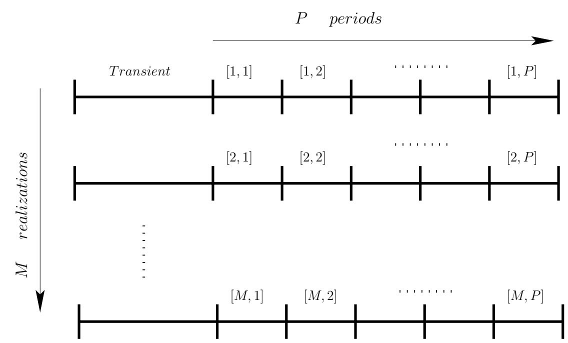

Fig. 3: Measurement scheme where P periods of the steady-state response to a random-phase multisine is meaured and repeated for M different random-phase realizations

Assuming there is no input noise, the nonparametric estimation of the BLA, the variance of the stochastic nonlinear distortions and the output noise variance, is based on the analysis of the sample mean and sample variance of the FRF over different multisine periods and phase realizations as shown in Fig.3. Since $N_y(jw_k)$ is a stochastic process and $Y_s(jw_k)$ is a periodic signal depending on the phase realization of the input signal, the FRF of the m^{th} phase realization and p^{th} period is related to the BLA, the stochastic nonlinear distortions and the output noise as

$$\begin{aligned} G^{[m,p]}(jw_k) &= \frac{Y^{[m,p]}}{U^{[m]}} \\ &= G_{BLA}(jw_k) + \frac{Y_S^{[m]}}{U^{[m]}} + \frac{N_Y^{[m,p]}}{U^{[m]}} \end{aligned} \tag{3}$$

This shows that the sample variance over the P periods only depends on the output noise, while the sample variance over the M realizations depends on both the stochastic nonlinear distortions and the output noise. From the

[1] Measurement scheme taken from [6] and written here for the sake of completeness

MxP noisy FRFs $G^{[m,p]}(jw_k)$, $m = 1,2,...,M$ and $p = 1,2,...,P$, one calculates for each multisine phase realization the average FRF, $\hat{G}^{[m]}(jw_k)$, and its sample variance, $\hat{\sigma}^2_{\hat{G}^{[m]}}(jw_k)$, over the P periods as

$$\hat{G}^{[m]}(jw_k) = \frac{1}{P}\sum_{p=1}^{P} G^{[m,p]}(jw_k) \tag{4}$$

$$\hat{\sigma}^2_{\hat{G}^{[m]}}(jw_k) = \sum_{p=1}^{P} \frac{|G^{[m,p]}(jw_k) - \hat{G}^{[m]}(jw_k)|^2}{P(P-1)} \tag{5}$$

Additional averaging over the M realizations gives the BLA $\hat{G}_{BLA}(jw_k)$ and the total variance $\hat{\sigma}^2_{\hat{G}_{BLA}}(jw_k)$ as

$$\hat{G}_{BLA}(jw_k) = \sum_{m=1}^{M} \frac{\hat{G}^{[m]}(jw_k)}{M} \tag{6}$$

$$\hat{\sigma}^2_{\hat{G}_{BLA}}(jw_k) = \sum_{m=1}^{M} \frac{|\hat{G}^{[m]}(jw_k) - \hat{G}_{BLA}(jw_k)|^2}{M(M-1)} \tag{7}$$

and an improved estimate of the noise variance

$$\hat{\sigma}^2_{\hat{G}_{BLA},n}(jw_k) = \frac{1}{M^2}\sum_{m=1}^{M} \hat{\sigma}^2_{\hat{G}^{[m]}}(jw_k) \tag{8}$$

Finally, the estimated variance of the stochastic nonlinear contributions is calculated as the difference between the total variance and the noise variance multiplied by the number of realizations (i.e., to account for the averaging process).

$$var(G_S(jw_k)) \approx M\left(\hat{\sigma}^2_{\hat{G}_{BLA}(jw_k)} - \hat{\sigma}^2_{\hat{G}_{BLA},n}(jw_k)\right) \tag{9}$$

2.3. Classification of nonlinearities

To classify the type of nonlinearities, odd multisines with a harmonic grid, so called odd-random multisines, are used. An odd multisine is a signal as in (2) where harmonics $k_i \in \{(2k-1) \mid k = 1,2,...,F\}$ (i.e., odd frequency lines) are excited. An odd-random multisine is an odd multisine where in every set of F_{group} consecutive odd harmonics one randomly chosen harmonic is not excited. The information about the stochastic nonlinear distortions is obtained via the detection lines (i.e. non-excited odd and even harmonics) in the output DFT spectrum. The system is said to have odd/even nonlinearity if the odd/even detection lines at the output DFT spectrum contain significant energy. The reason is that the response of an odd/even nonlinearity to odd harmonics generates odd/even harmonics only [6].

Assuming that the input detection lines have zero magnitude, the presence of signal energy at the corresponding output detection lines is due to the nonlinear response of the system plus some output noise. The stochastic nonlinear contributions and noise estimations are based on the analysis of the sample mean and variance of the output spectra. The measurement scheme is the same as in Fig.3 with $M = 1$ but now with an increased number of periods.

The average output spectrum is calculated at all frequencies (excited and non-excited harmonics) as [6]

$$\hat{Y}(jw_k) = \frac{1}{P}\sum_{p=1}^{P} Y^{[p]}(jw_k) \tag{10}$$

3. MEASUREMENTS

Measurements were done in a room that is acoustically conditioned and prepared to have low reverberation times and listening comfort but that is not anechoic. Hence it is expected that some (linear) dynamics are added to the frequency responses. The microphone is a high-quality condenser microphone connected through a mixing console to a PC via a high-quality sound card. It is expected that none of these elements will cause any significant harmonic distortion.

Three different active loudspeakers were measured following the schemes of section 2: One desktop PC loudspeaker, one medium-size loudspeaker typically used in small teleconferencing rooms and one high-quality recording studio monitor. The quality of the loudspeakers and their usage differ greatly. The input signal frequency range is $f_0 = 1$ Hz, $f_{max} = 10$ kHz with $k_1 = 250$ and relative input power RMS = 0.02, 0.04, 0.08 and 0.16 for the exciting multisines.

3.1. Estimating the level of the stochastic nonlinear contributions

The level of the stochastic nonlinear contribution ($var(G_S(jw_k))$) is calculated for different RMS values of a full multisine as explained in section 2.2 with $P = 2$ and $M = 7$. In the following figures the upper line ($-.-$) represent the $\hat{G}_{BLA}$, stars ($\star$) represent the stochastic nonlinear contribution and dots (.) represent the output noise variance. It can be seen from Fig.4 and 5 that for the PC and teleconferencing loudspeakers the level of the stochastic nonlinear contribution increases steadily with increasing input power reaching significant levels quite easily. On the other hand, nonlinear distortion in the monitor loudspeaker remains at low levels (50 dB below $\hat{G}_{BLA}$) for every input RMS value (see Fig.6). The gain of the studio monitor was further increased and the measuring process repeated. Results are given in Fig.7. It can be seen that the level of the nonlinear source is still low until a certain input level is reached, where the nonlinear source level increases significantly. This suggests that due to the high quality of the loudspeaker, the nonlinear distortion is kept small (50 dB) even for high amplitudes. However, this loudspeaker has a protection device (i.e. saturation curve) in the amplifier to avoid diaphragm damage at too high input signals. This saturation curve is what causes significant nonlinear distortion at very high amplitudes.

It is clear from this section that nonlinearities are present in low-quality loudspeakers and eventually in high-quality loudspeakers as well.

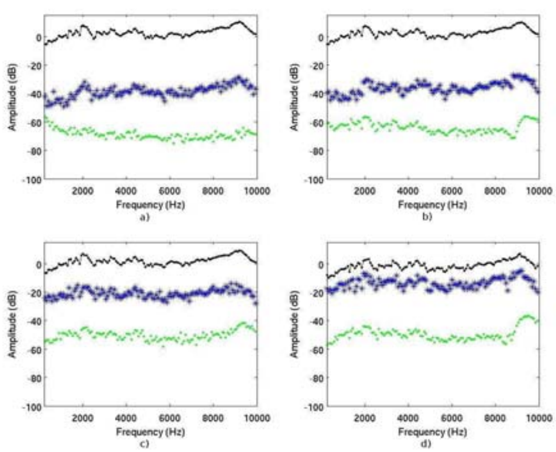

Fig. 4: $\hat{G}_{BLA}$, nonlinear distortion and noise level of the PC loudspeaker at RMS input signal a) 0.02, b) 0.04, c) 0.08, d) 0.16

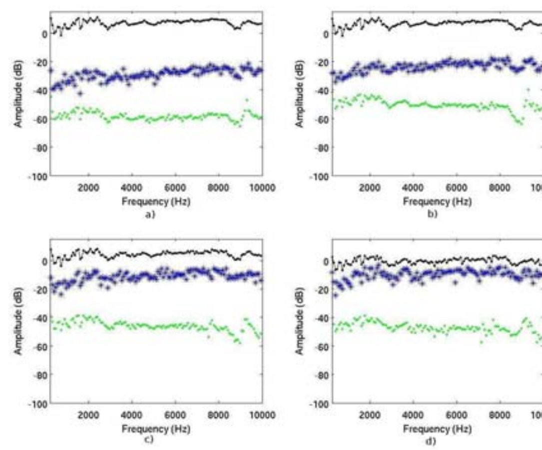

Fig. 5: $\hat{G}_{BLA}$, nonlinear distortion and noise level of the teleconferencing loudspeaker at RMS input signal a) 0.02, b) 0.04, c) 0.08, d) 0.16

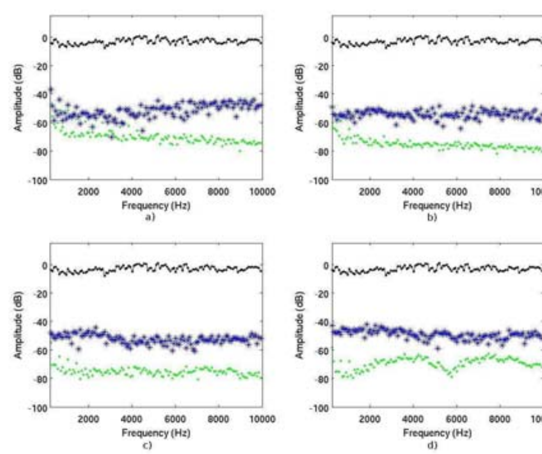

Fig. 6: $\hat{G}_{BLA}$, nonlinear distortion and noise level of the studio monitor at RMS input signal a) 0.02, b) 0.04, c) 0.08, d) 0.16

3.2. Classifying nonlinearities

To classify the nonlinear distortion in odd and even nonlinearities, an odd-random multisine is created. The output is measured following the scheme of section 2.3, with $M = 1$ and $P = 12$. The frequency range is the same as in 3.1, but now only odd harmonics are excited, with $F_{group} = 3$. By inspecting the odd and even detection lines in the output DFT spectrum we can see the odd and even nonlinear response of the loudspeaker. In the figures the upper line ($-.-$) is the output level at the excited lines, circles ($\bigcirc$) represent the output level at non-excited odd harmonics and stars ($\star$) represent the level at non-excited even harmonics. From Fig.8, 9 and 10 it can be seen that the predominant type of nonlinearities are odd. The lower-quality loudspeakers (i.e., PC and teleconferencing) generate odd harmonics that increase in amplitude

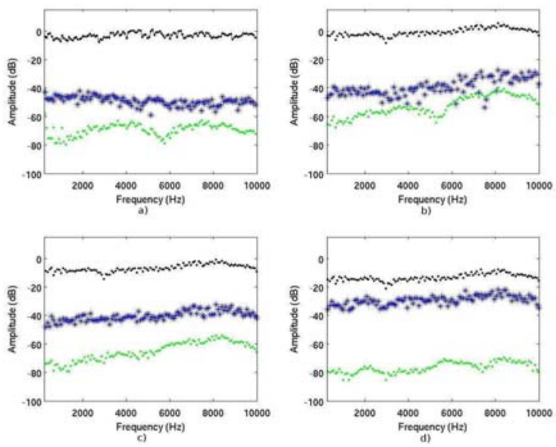

Fig. 7: $\hat{G}_{BLA}$, nonlinear distortion and noise level of the studio monitor with increased gain at RMS input signal a) 0.02, b) 0.04, c) 0.08, d) 0.16

with increasing input level. Moreover, the output level (–.–) increases with input level correspondingly. On the other hand, the output level (–.–) of the high quality loudspeaker stops increasing in Fig.10(d) while the odd nonlinear distortion does not. This is again a consequence of the amplifier saturation curve.

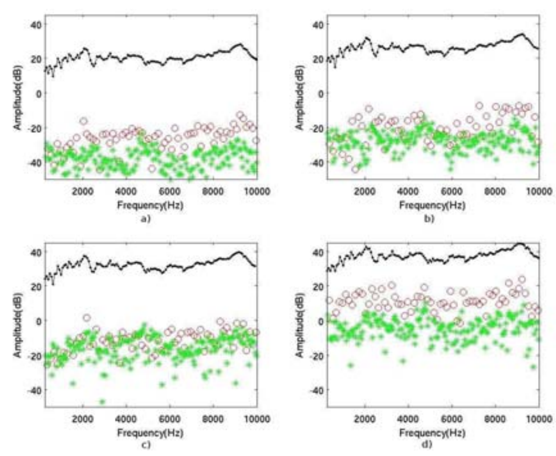

Fig. 8: Output spectrum of the PC loudspeaker at RMS input signal a) 0.02, b) 0.04, c) 0.08, d) 0.16

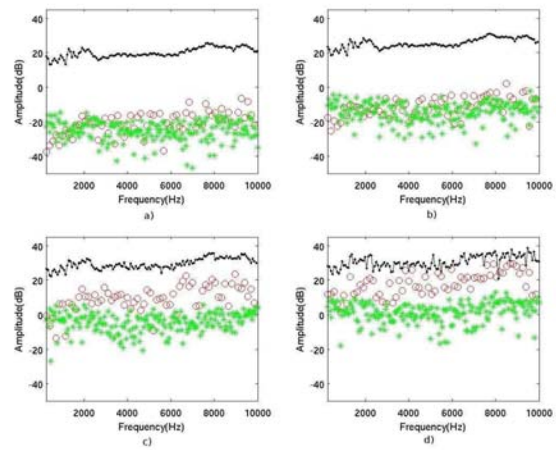

Fig. 9: Output spectrum of the teleconferencing loudspeaker at RMS input signal a) 0.02, b) 0.04, c) 0.08, d) 0.16

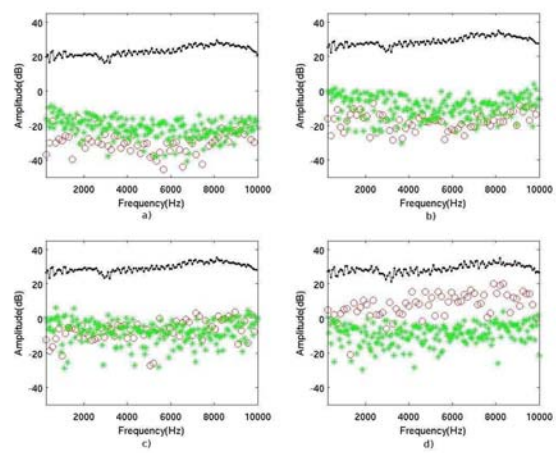

Fig. 10: Output spectrum of the studio monitor at RMS input signal a) 0.02, b) 0.04, c) 0.08, d) 0.16

3.3. Discussion

It has been shown that at relatively high input signal levels, nonlinear distortions should be taken into account. In this case odd nonlinearities are significantly more predominant than even nonlinearities in the full frequency range. This implies that the dominant nonlinearity must be modelled as a third-order (cubic) system. The critical problem is that, in Volterra filters, the number of parameters increases exponentially with nonlinear order. Therefore, second-order Volterra filters are typically used in AEC applications. The problem now becomes obvious, if second-order Volterra filters are computationally very intensive, third-order Volterra filters are impractical in any AEC application.

In [4] and [7] it is shown that a physical model based on a nonlinear version of the well-known lumped parameter model of the loudspeaker can achieve a good representation of the odd harmonics when comparing simulations with measurements. This modelling approach is based on the dependence of the lumped parameters with the voice coil position. It is generally accepted that the variations of the magnetic force factor, the voice coil inductance and compliance of the suspension, are the main causes of nonlinear distortion. These are modelled as functions of cone displacement and therefore they are static (memoryless) elements in the model. This implies that the odd nonlinear distortions should be modelled as being memoryless. This is also in conflict with the general use of Volterra filters with memory to model loudspeaker nonlinearities [1].

4. CONCLUSION

A conclusion drawn from the measurements is that the level of nonlinearities in all measured loudspeakers is significant. The level of the nonlinearities depends on the quality of the (active) loudspeaker, although eventually any (active) loudspeaker suffers from significant nonlinear distortion. In AEC applications, nonlinear distortions are of great importance if the levels are high. As opposed to the situation when only background noise is present (i.e., where increasing the level of the loudspeaker signal will increase the SNR of the echo signal; hence improve the performance of the AEC), if the nonlinear source is present, increasing the level of the input signal will also increase the level of the nonlinear noise source.

Therefore nonlinear adaptive filters are needed in AEC applications to account for nonlinear distortions. In the literature second-order Volterra filters with memory are generally used; however they can model only second-order even nonlinearities. Measurements show that odd nonlinearities are indeed predominant over the entire frequency range. This leads to the second conclusion that at least third-order filters should be applied to account for odd nonlinearities.

The use of third-order adaptive Volterra filters could become prohibitive in AEC applications, however it is shown that loudspeaker odd nonlinearities can be modelled as static functions of the voice coil position. Hence, the third conclusion and comment on further research is that the modelling effort in AEC application should be put in memoryless adaptive filters featuring a block structure (i.e. static nonlinearities and linear dynamic blocks combined), which reduces the number of parameters to be estimated.

5. ACKNOWLEDGEMENTS

This research work was carried out at the ESAT laboratory of Katholieke Universiteit Leuven, in the frame of the EST-SIGNAL Marie-Curie Fellowship program (http://est-signal.i3s.unice.fr) under contract No. MEST-CT-2005-021175, and the Concerted Research Action GOA-AMBioRICS.

The authors wish to thank Prof. Johan Schoukens and co-workers for clarifying explanations, instructive discussions and general support which were essential for writing this paper.

6. REFERENCES

[1] A. Stenger and R. Rabenstein, ”Adaptive Volterra filters for acoustic echo cancellation,” *in Proc. IEEE Workshop Nonlinear Signal Image Process.* Antalya, Turkey, pp. 679-683, June, 1999.

[2] F. Kuech and W. Kellerman, ”Partitioned block frequency-domain adaptive second-order Volterra filter”, *IEEE Trans. Signal Process.*, vol. 53, no. 2, Feb. 2005, pp 564-575.

[3] S. Haykin. *Adaptive Filter Theory*. New Jersey: Prentice Hall, 1996.

[4] B. R. Pedersen. *Error correction of loudspeakers*. PhD thesis. Aalborg University, Denmark. May 2008

[5] R. Pintelon and J. Schoukens *System Identification: A frequency domain approach*. New York: Wiley- IEEE Press. May 2001.

[6] R. Pintelon, G. Vandersteen, L. De Locht, Y. Rolain, and J. Schoukens, ”Experimental characterization of operational amplifiers: a system identification approach - Part I: theory and simulations”, *IEEE Trans. Instrum. Meas.*, vol. 53, no. 3, pp. 854-862, June 2004.

[7] F.T. Agerkvist, ”Modelling loudspeaker non-linearities”. Proc. AES 32^{nd} AES Conf. DSP for loudspeakers, Hillerød Denmark, Sept. 2007, paper no. 7.